Cross-lingual Projection of Role-Semantic Information

Sebastian Padó

Computational Lingustics
Saarland University
pado@coli.uni-sb.de

October 28, 2005

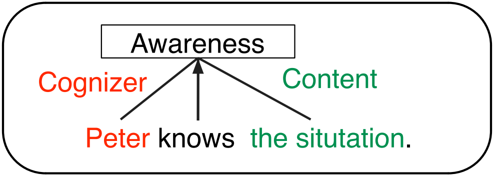
Outline

1. Motivation
   - Shallow Semantic Parsing
   - Knowledge Acquisition Bottleneck
2. Role Projection in a Parallel Corpus
   - Word-based Projection
   - Syntax-based Projection
3. Projection Results
   - Experimental Set-up
   - Evaluation of Projection Models

Outline

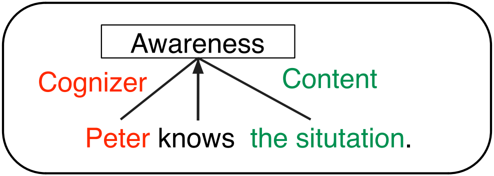
Shallow Semantic Parsing

The task of automatically identifying the **semantic roles** conveyed by sentential constituents.



Shallow Semantic Parsing

The task of automatically identifying the **semantic roles** conveyed by sentential constituents.



- Relevant for several applications (IE, IR, QA)
- Common semantic representation across languages

Frame Semantics

Role-semantics paradigm based on **conceptual** structures (Filmore et al., 2003).

	Frame: AWARENESS	
Frame Elements	COGNIZER	**Peter** knows the situation. **Pat** believes that things will change.
	CONTENT	Peter knows **the situation.** Pat believes **that things will change.**
FEEs	aware.v, believe.v, comprehend.v, conceive.v, imagine.v, know.v, belief.n, consciousness.v, hunch.n, suspicion.v, conscious.a, knowledgeable.a	

Outline

1. Motivation
   - Shallow Semantic Parsing
   - Knowledge Acquisition Bottleneck
2. Role Projection in a Parallel Corpus
   - Word-based Projection
   - Syntax-based Projection
3. Projection Results
   - Experimental Set-up
   - Evaluation of Projection Models

Knowledge Acquisition Bottleneck

- Data-driven development of shallow semantic parsers (see e.g. Carreras and Màrquez, 2005) requires:
  1. English FrameNet lexicon ($>$ 500 frames, $>$ 7,000 lemmas)
  2. English annotated example sentences (100,000 available)

Knowledge Acquisition Bottleneck

- Data-driven development of shallow semantic parsers (see e.g. Carreras and Màrquez, 2005) requires:
  1. English FrameNet lexicon ($>$ 500 frames, $>$ $7,000$ lemmas)
  2. English annotated example sentences (100,000 available)

- Frame Semantics is (largely) language-independent: annotation efforts for German, Spanish, and Japanese
- Annotation laborious and time-consuming

Knowledge Acquisition Bottleneck

- Data-driven development of shallow semantic parsers (see e.g. Carreras and Màrquez, 2005) requires:
  1. English FrameNet lexicon ($>$ 500 frames, $>$ $7,000$ lemmas)
  2. English annotated example sentences (100,000 available)
- Frame Semantics is (largely) language-independent: annotation efforts for German, Spanish, and Japanese
- Annotation laborious and time-consuming

Knowledge Acquisition Bottleneck:
Can we reduce annotation effort for new languages?

Main Ideas

- Use English FrameNet resource as basis
- Project information to other languages using parallel corpora

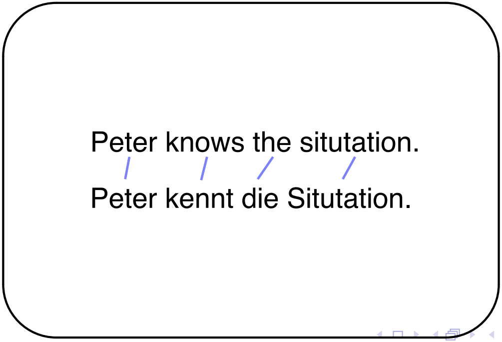
Two steps:

1. Project FrameNet lexicon (IGK meeting in Mertesdorf)
2. Project role information (now)

Outline

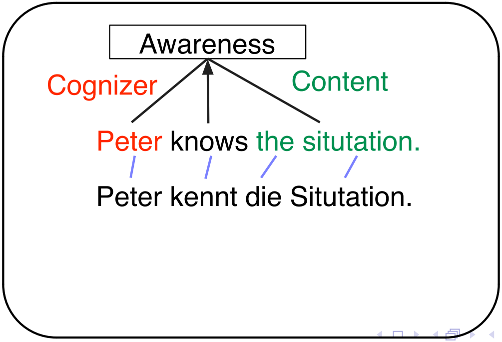
Role Projection in a Parallel Corpus

1. Start with bi-sentence (translation) with word alignment



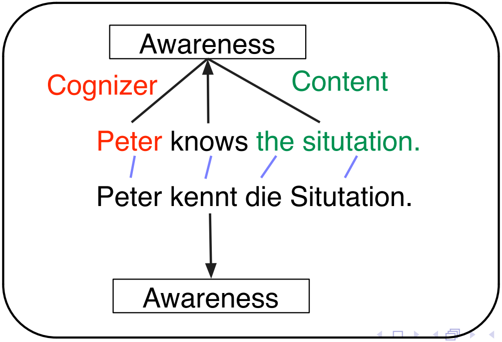
Role Projection in a Parallel Corpus

1. Start with bi-sentence (translation) with word alignment
2. Obtain role-semantic analysis for source sentence



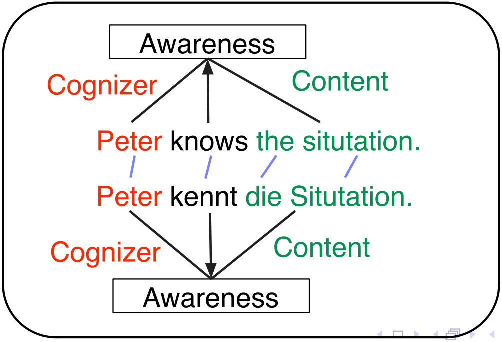
Role Projection in a Parallel Corpus

1. Start with bi-sentence (translation) with word alignment
2. Obtain role-semantic analysis for source sentence
3. Check if target predicate can evoke the same frame



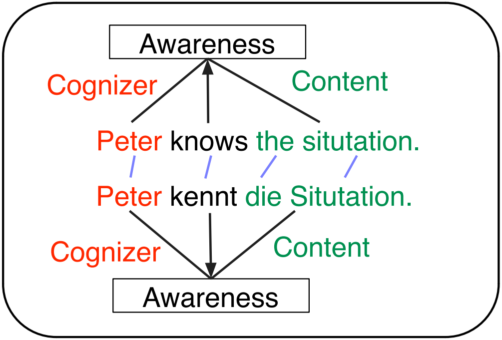
Role Projection in a Parallel Corpus

1. Start with bi-sentence (translation) with word alignment
2. Obtain role-semantic analysis for source sentence
3. Check if target predicate can evoke the same frame
4. Project roles from source to target sentence



Role Projection in a Parallel Corpus

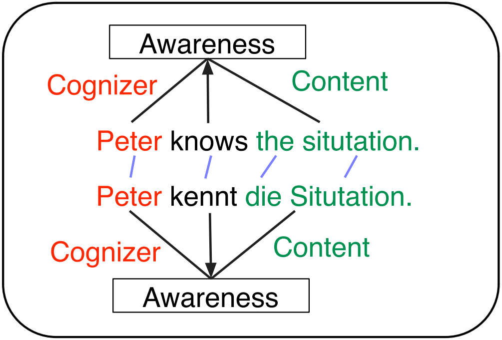
1. Start with bi-sentence (translation) with word alignment
2. Obtain role-semantic analysis for source sentence
3. Check if target predicate can evoke the same frame
4. Project roles from source to target sentence



Assumption:
Bi-sentences have parallel (role) semantics

Role Projection in a Parallel Corpus

1. Start with bi-sentence (translation) with word alignment
2. Obtain role-semantic analysis for source sentence
3. Check if target predicate can evoke the same frame
4. Project roles from source to target sentence



Assumption:
Bi-sentences have parallel (role) semantics
Empirical result:
For English / German, **92%** of roles match

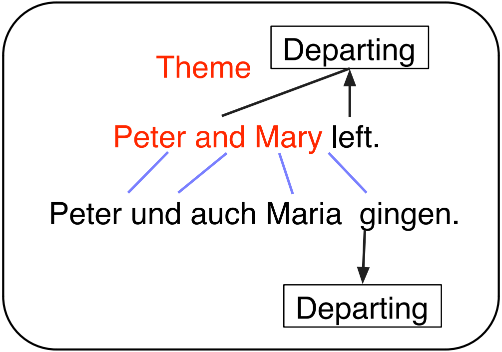
Related Work

- Induction of multilingual morphological analyzers (Mann and Yarowsky, 2001)
- Projection of POS-tag information (Yarowsky et al., 2001)
- Projection of bracketing information (Yarowsky et al., 2001)
- Projection of dependency relations (Hwa et al., 2002)

Outline

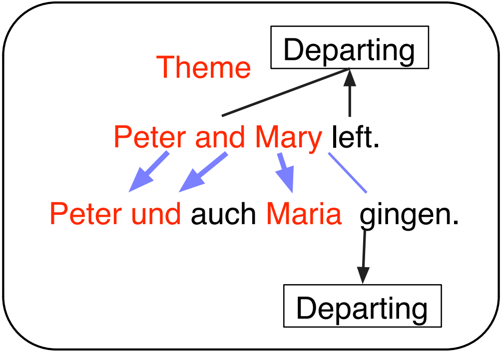
Word-based Projection

1. For each source semantic role, identify word span



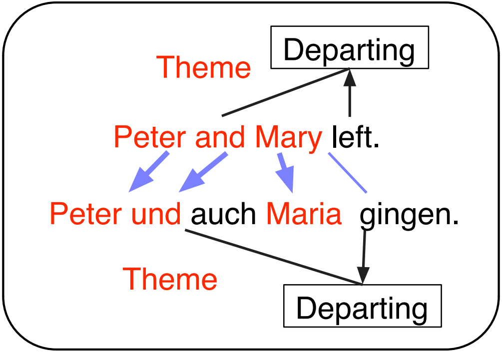
Word-based Projection

1. For each source semantic role, identify word span
2. Follow all word alignments



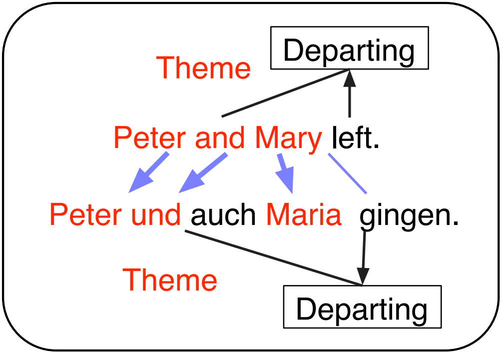
Word-based Projection

1. For each source semantic role, identify word span
2. Follow all word alignments
3. Target role span is union of all projections



Word-based Projection

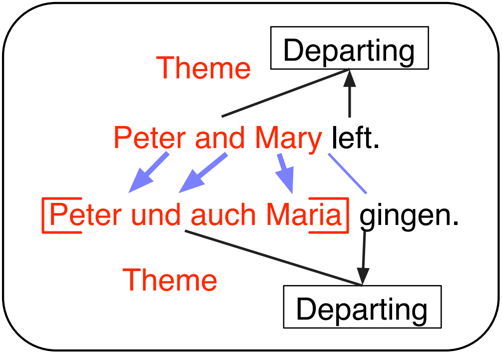
1. For each source semantic role, identify word span
2. Follow all word alignments
3. Target role span is union of all projections



Missing word alignments:
convex complementing heuristic

Word-based Projection

1. For each source semantic role, identify word span
2. Follow all word alignments
3. Target role span is union of all projections



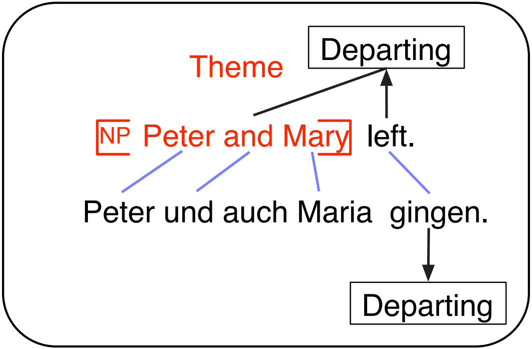
Missing word alignments: **convex complementing heuristic**

Outline

1. Motivation
   - Shallow Semantic Parsing
   - Knowledge Acquisition Bottleneck
2. Role Projection in a Parallel Corpus
   - Word-based Projection
   - Syntax-based Projection
3. Projection Results
   - Experimental Set-up
   - Evaluation of Projection Models

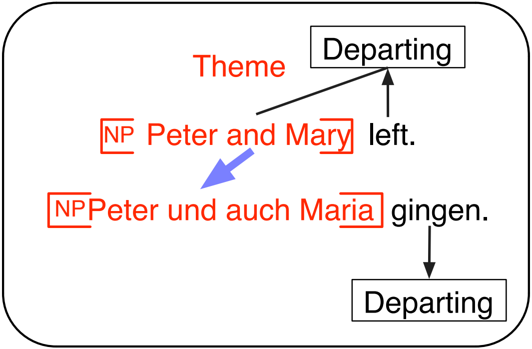
Syntax-based Projection

1. For each source role, identify source constituent(s)



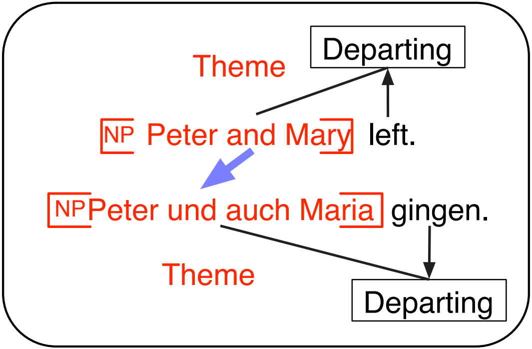
Syntax-based Projection

1. For each source role, identify source constituent(s)
2. Find **optimal alignment** between S and T constituents



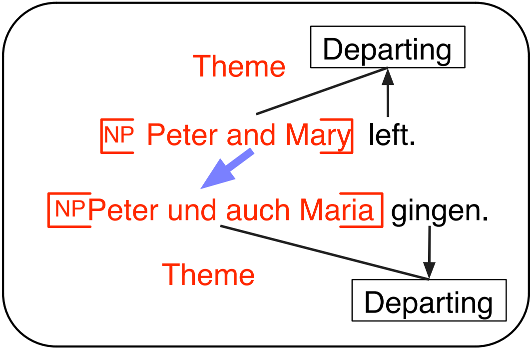
Syntax-based Projection

1. For each source role, identify source constituent(s)
2. Find **optimal alignment** between S and T constituents
3. Label target constituent(s) with role



Syntax-based Projection

1. For each source role, identify source constituent(s)
2. Find **optimal alignment** between S and T constituents
3. Label target constituent(s) with role



Can use any bracketing information (chunks, "full" constituents)

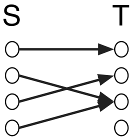
Probabilistic Constituent Alignment

- Two sets of constituents, C and C'
- For each $c \in C$, find $c' \in C'$ with maximal word overlap

Probabilistic Constituent Alignment

- Two sets of constituents, C and C'
- For each $c \in C$, find $c' \in C'$ with maximal word overlap

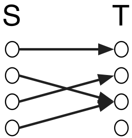
1. **Forward** alignment
   - Align from source to target constituents
   - Assumes one target constituent per source constituent



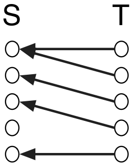
Probabilistic Constituent Alignment

- Two sets of constituents, C and C'
- For each $c \in C$, find $c' \in C'$ with maximal word overlap

1. **Forward** alignment
   - Align from source to target constituents
   - Assumes one target constituent per source constituent



2. **Backward** alignment
   - Aligns from target to source constituents
   - Source constituents can correspond to none or $>$ 1 target constituents



Outline

1. Motivation
   - Shallow Semantic Parsing
   - Knowledge Acquisition Bottleneck
2. Role Projection in a Parallel Corpus
   - Word-based Projection
   - Syntax-based Projection
3. Projection Results
   - Experimental Set-up
   - Evaluation of Projection Models

Experimental Set-up

Data

- Sample of 1000 English-German Bi-sentences from EUROPARL (Koehn, 2000)
- Choice informed by FrameNet (E) and SALSA (D) lexicons
- Two sides of each bi-sentence annotated independently
  - Annotators tagged equal amount of English and German
  - Inter-annotator agreement: $\kappa = 0.84$
- Word alignment: GIZA++ (Och and Ney, 2003)

Experimental Set-up

Method

- Project roles from English gold annotation onto German
- Evaluate against German gold annotation
- Compare word-based, chunk-based, and constituent-based models
  - Chunk-based models use Abney's (1997, E) and Schmid and Schulte im Walde's (2000, D) base NP chunkers
  - Constituent-based models use Collins' (1997, E) and Dubey's (2003, D) parsers

Outline

Word-based Projection

Model	Precision	Recall	F-score
WordAlign	0.41	0.40	0.41
WordAlign + ConvexComp	**0.46**	**0.45**	**0.46**
UpperBnd	0.85	0.84	0.84

- Alignments can guide projection task substantially
    - WordAlign exploits no linguistic information
- ConvexComp improves the F-score
- F-score sig worse than UpperBnd

Chunk-based Projection

Model	Prec	Recall	F-score
WordAlign + ConvexComp	0.46	0.45	0.46
ForwardAlign	**0.46**	**0.25**	**0.32**
BackwardAlign	0.30	0.24	0.27
BackwardAlign + ConvexComp	0.32	0.26	0.29
UpperBnd	0.85	0.84	0.84

- ForwardAlign sig worse Recall than WordAlign
  - F-score sig worse than UpperBnd
- Problem: Often, no chunks for source **and** target role span
  - Overlap maximisation does not yield sensible results

Syntax-based Projection

Model	Prec.	Recall	F-score
WordAlign + ConvexComp	0.46	0.45	0.46
ForwardAlign	**0.70**	**0.60**	**0.65**
BackwardAlign	0.60	0.46	0.52
BackwardAlign + ConvexComp	0.74	0.56	0.64
UpperBnd	0.85	0.84	0.84

- ForwardAlign sig better than WordAlign and BackwardAlign; sig worse than UpperBnd
- One-to-one assumption (ForwardAlign) mostly warranted
- Less guided alignment (BackwardAlign) requires ConvexComp

Error Analysis

1. Wrong or missing word alignments (e.g., PPs)

   He asks [MSG for a doctor].

   Er fragt nach einem Arzt.

2. Translational divergences (wrong or missing projections)

   We claim and [SPKR we] say [MSG ...]

   [SPKR Wir] behaupten und -- sagen [MSG ...]

Conclusions

Summary

- Principled framework for role projection
- Semantic roles can be projected between languages
- Bracketing can make up for problems in word alignment
- Best model performs at 0.65 F-Score (UpperBnd is 0.84)
- Base NP chunks not sufficient